

# Spectral Compressed Sensing via CANDECOMP/PARAFAC Decomposition of Incomplete Tensors

Jun Fang, Linxiao Yang, and Hongbin Li, *Senior Member, IEEE*

**Abstract**—We consider the line spectral estimation problem which aims to recover a mixture of complex sinusoids from a small number of randomly observed time domain samples. Compressed sensing methods formulates line spectral estimation as a sparse signal recovery problem by discretizing the continuous frequency parameter space into a finite set of grid points. Discretization, however, inevitably incurs errors and leads to deteriorated estimation performance. In this paper, we propose a new method which leverages recent advances in tensor decomposition. Specifically, we organize the observed data into a structured tensor and cast line spectral estimation as a CANDECOMP/PARAFAC (CP) decomposition problem with missing entries. The uniqueness of the CP decomposition allows the frequency components to be super-resolved with infinite precision. Simulation results show that the proposed method provides a competitive estimate accuracy compared with existing state-of-the-art algorithms.

**Index Terms**—CANDECOMP/PARAFAC decomposition, line spectral estimation, super resolution.

## I. INTRODUCTION

The problem of recovering the frequency components of a mixture of complex sinusoids from a finite number of time samples arises in a variety of applications, such as radar, sonar, array signal processing and seismology. Such a problem has been extensively investigated over the past decades and a number of classical methods such as the MUSIC [1], ESPRIT [2], matrix-pencil [3], and many others were proposed as early as 1980s. In these studies, the shift invariance property of harmonic structures, i.e. the subspace of a consecutive segment of time domain samples remains unaltered irrespective of the starting point of the segment, was usually exploited for algorithm development. These methods, however, require that the sampling rate satisfies the Nyquist-Shannon sampling theorem.

Another line of work that has attracted much attention recently is to formulate line spectral estimation as a sparse signal recovery (i.e. compressed sensing) problem. By exploiting the sparsity in the frequency domain, compressed

sensing techniques allow the frequency components to be recovered from only a small, random subset of uniformly spaced samples. The sampling rate can thus be significantly reduced. Also, unlike classical methods [1]–[3] which assume the knowledge of the number of frequency components, compressed sensing methods are able to determine the model order in an automatic manner. Nevertheless, to apply the compressed sensing technique to the line spectral estimation problem, one has to discretize the continuous parameter space into a finite set of grid points and assumes that the true parameters lie on the discretized grid. Grid mismatch arises when the true parameters are inconsistent with the discretized grid, in which case compressed sensing methods may incur a considerable performance degradation. To address this issue, a class of off-grid (or super-resolution) compressed sensing approaches were proposed, e.g. [4]–[8]. Specifically, in [6], [7], an atomic norm-minimization approach was proposed to handle the infinite dictionary with continuous atoms. It was shown that given that the frequency components are sufficiently separated, the frequency components of a mixture of complex sinusoids can be super-resolved with infinite precision from only coarse-scale measurements. Also, in [9], [10], by arranging the observed samples into a low-rank Hankel matrix, a structured matrix completion method was developed to recover the real-valued frequency parameters.

In this paper, we propose a new method which organizes the observed data into a structured tensor and cast line spectral estimation as a CANDECOMP/PARAFAC (CP) decomposition problem with missing entries. Due to the uniqueness of the CP decomposition, the frequency components of a mixture of complex sinusoids is guaranteed to be super-resolved from only a small number of nonuniform samples. Simulation results show that the proposed method provides competitive estimation performance compared with existing state-of-the-art algorithms.

## II. PRELIMINARIES

We first provide a brief review on tensor and the CP decomposition. A tensor is the generalization of a matrix to higher order dimensions, also known as ways or modes. Vectors and matrices can be viewed as special cases of tensors with one and two modes, respectively.

Let  $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$  denote an  $N$ th order tensor with its  $(i_1, \dots, i_N)$ th entry denoted by  $\mathcal{X}_{i_1, \dots, i_N}$ . Here the order  $N$  of a tensor is the number of dimensions. Fibers are the higher-order

Jun Fang, and Linxiao Yang are with the National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China, Email: Jun-Fang@uestc.edu.cn

Hongbin Li is with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030, USA, E-mail: Hongbin.Li@stevens.edu

This work was supported in part by the National Science Foundation of China under Grant 61172114, and the National Science Foundation under Grant ECCS-1408182.

analogue of matrix rows and columns. The mode- $n$  fibers of  $\mathcal{X}$  are  $I_n$ -dimensional vectors obtained by fixing every index but  $i_n$ . Unfolding or matricization is an operation that turns a tensor to a matrix. Specifically, the mode- $n$  unfolding of a tensor  $\mathcal{X}$ , denoted as  $\mathbf{X}_{(n)}$ , arranges the mode- $n$  fibers to be the columns of the resulting matrix. The  $n$ -mode product of  $\mathcal{X}$  with a matrix  $\mathbf{A} \in \mathbb{R}^{J \times I_n}$  is denoted by  $\mathcal{X} \times_n \mathbf{A}$  and is of size  $I_1 \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N$ , with each mode- $n$  fiber multiplied by the matrix  $\mathbf{A}$ , i.e.

$$\mathcal{Y} = \mathcal{X} \times_n \mathbf{A} \Leftrightarrow \mathbf{Y}_{(n)} = \mathbf{A} \mathbf{X}_{(n)} \quad (1)$$

The CP decomposition decomposes a tensor into a sum of component rank-one component tensors, i.e.

$$\mathcal{X} = \sum_{r=1}^R \lambda_r \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \cdots \circ \mathbf{a}_r^{(N)} \quad (2)$$

where  $\mathbf{a}_r^{(n)} \in \mathbb{R}^{I_n}$ , ‘ $\circ$ ’ denotes the vector outer product, the minimum achievable  $R$  is referred to as the rank of the tensor, and  $\mathbf{A}^{(n)} \triangleq [\mathbf{a}_1^{(n)} \cdots \mathbf{a}_{R_n}^{(n)}] \in \mathbb{R}^{I_n \times R_n}$  denotes the factor matrix along the  $n$ -th mode. Elementwise, we have

$$\mathcal{X}_{i_1 i_2 \dots i_N} = \sum_{r=1}^R \lambda_r a_{i_1 r}^{(1)} a_{i_2 r}^{(2)} \cdots a_{i_N r}^{(N)} \quad (3)$$

### III. TENSOR FORMULATION

Consider the line spectral estimation problem where the signal  $x_n$  can be represented as a summation of a number of complex sinusoids, i.e.

$$x_n = \sum_{k=1}^K a_k e^{-j\omega_k(n-1)}, \quad n = 1, 2, \dots, N \quad (4)$$

where  $\omega_k \in [0, 2\pi)$  and  $a_k$  denote the frequency and complex amplitude of the  $k$ -th component, respectively. Let  $\mathbf{y} \triangleq [y_1 \ y_2 \ \dots \ y_M]^T$  denote the noise-corrupted observations randomly chosen from the original set  $\{x_n\}$ . Our objective is to estimate  $\{a_k\}$  and  $\{\omega_k\}$  from the observed data  $\mathbf{y}$ . To this objective, we organize the observed samples  $\{y_m\}_{m=1}^M$  as an incomplete structured third-order tensor. The unknown parameters  $\{\omega_k\}$  along with the missing entries can be estimated via the CP decomposition of this incomplete tensor.

To better illustrate our method, we first show how to construct the tensor using the original samples  $\{x_n\}$  instead of the observed samples  $\{y_n\}$ . In order to obtain a tensor which admits a CP decomposition, we arrange samples  $\{x_n\}$  to form a third-order tensor  $\mathcal{X} \in \mathbb{C}^{(N-L-P+2) \times L \times P}$ , with each slice along the third-mode of the tensor being an  $(N-L-P+2) \times L$  matrix, i.e.

$$\mathcal{X}(:, :, i) = [\mathbf{x}_{L+i-1} \ \mathbf{x}_{L+i-2} \ \dots \ \mathbf{x}_i] \quad (5)$$

where  $L$  and  $P$  are parameters whose choice will be discussed later, and

$$\mathbf{x}_t \triangleq [x_t \ x_{t+1} \ \dots \ x_{N-L-P+t+1}]^T \quad t = 1, \dots, L+P-1 \quad (6)$$

By exploiting the inherent structure, each slice  $\mathcal{X}(:, :, i)$  can be expressed as

$$\mathcal{X}(:, :, i) = \mathbf{A} \mathbf{D}_i \mathbf{B}^T \quad i = 1, \dots, P \quad (7)$$

where  $\mathbf{A} \in \mathbb{C}^{(N-L-P+2) \times K}$ ,  $\mathbf{B} \in \mathbb{C}^{L \times K}$  and  $\mathbf{D}_i \in \mathbb{C}^{K \times K}$  are defined respectively as

$$\mathbf{A} \triangleq \begin{bmatrix} 1 & \cdots & 1 \\ e^{-j\omega_1} & \cdots & e^{-j\omega_K} \\ \vdots & & \vdots \\ e^{-j\omega_1(N-L-P+1)} & \cdots & e^{-j\omega_K(N-L-P+1)} \end{bmatrix} \quad (8)$$

$$\mathbf{B} \triangleq \begin{bmatrix} e^{-j\omega_1(L-1)} & \cdots & e^{-j\omega_K(L-1)} \\ e^{-j\omega_1(L-2)} & \cdots & e^{-j\omega_K(L-2)} \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix} \quad (9)$$

and

$$\mathbf{D}_i \triangleq \text{diag}(a_1 e^{-j\omega_1(i-1)}, a_2 e^{-j\omega_2(i-1)}, \dots, a_K e^{-j\omega_K(i-1)}) \quad (10)$$

Let  $\mathbf{a}_k$  and  $\mathbf{b}_k$  denote the  $k$ th column of  $\mathbf{A}$  and  $\mathbf{B}$ , respectively. The slice  $\mathcal{X}(:, :, i)$  can be rewritten as

$$\mathcal{X}(:, :, i) = \sum_{k=1}^K a_k e^{-j\omega_k(i-1)} \mathbf{a}_k \mathbf{b}_k^T \quad (11)$$

We see that each slice of  $\mathcal{X}$  is a weighted sum of a common set of rank-one outer products. Hence the tensor  $\mathcal{X}$  admits the following CP decomposition which decomposes a tensor as a sum of component rank-one tensors, i.e.

$$\mathcal{X} = \sum_{k=1}^K \mathbf{a}_k \circ \mathbf{b}_k \circ \mathbf{c}_k \quad (12)$$

where  $\mathbf{c}_k \triangleq [a_k \ a_k e^{-j\omega_k} \ \dots \ a_k e^{-j\omega_k(M-1)}]^T$ . Define  $\mathbf{C} \triangleq [\mathbf{c}_1 \ \dots \ \mathbf{c}_K]$ . The matrices  $\mathbf{A}$ ,  $\mathbf{B}$  and  $\mathbf{C}$  are factor matrices associate with the tensor  $\mathcal{X}$ . Since  $K$  is usually small, the above factorization implies that the tensor  $\mathcal{X}$  has a low-rank structure.

### IV. ALGORITHM DEVELOPMENT

When only the observations  $\{y_m\}$ , are available, we can readily construct an incomplete third-order tensor  $\mathcal{Y}$  by following the way we construct  $\mathcal{X}$ . By exploiting the low rank structure, the missing entries of  $\mathcal{Y}$ , along with the factor matrices that contain information about the frequencies, can be estimated. Specifically, the problem can be cast as

$$\begin{aligned} \min_{\mathcal{X}} \quad & \text{rank}(\mathcal{X}) \\ \text{s.t.} \quad & \|\mathcal{O} * \mathcal{Y} - \mathcal{O} * \mathcal{X}\|_F^2 \leq \varepsilon \end{aligned} \quad (13)$$

where  $\mathcal{O}$  is a binary tensor of the same size as  $\mathcal{X}$  with  $\mathcal{O}_{ijk} = 1$  if  $\mathcal{X}_{ijk}$  is observed, and  $\mathcal{O}_{ijk} = 0$  otherwise.  $\varepsilon$  is an error tolerance parameter related to noise statistics. Note that the CP rank is the minimum number of rank-one tensor components required to represent the tensor. Thus the search of a low rank

$\mathcal{X}$  can be converted to optimization of its associated factor matrices. Let

$$\mathcal{X} = \sum_{k=1}^{\tilde{K}} \tilde{\mathbf{a}}_k \circ \tilde{\mathbf{b}}_k \circ \tilde{\mathbf{c}}_k \quad (14)$$

where  $\tilde{K} \gg K$ , and

$$\begin{aligned} \tilde{\mathbf{A}} &\triangleq [\tilde{\mathbf{a}}_1 \ \dots \ \tilde{\mathbf{a}}_{\tilde{K}}] \\ \tilde{\mathbf{B}} &\triangleq [\tilde{\mathbf{b}}_1 \ \dots \ \tilde{\mathbf{b}}_{\tilde{K}}] \\ \tilde{\mathbf{C}} &\triangleq [\tilde{\mathbf{c}}_1 \ \dots \ \tilde{\mathbf{c}}_{\tilde{K}}] \end{aligned}$$

The optimization (13) can be re-expressed as

$$\begin{aligned} \min_{\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}}} \quad & \|\mathbf{z}\|_0 \\ \text{s.t.} \quad & \|\mathcal{O} * \mathbf{y} - \mathcal{O} * \mathcal{X}\|_F^2 \leq \varepsilon \\ & \mathcal{X} = \sum_{k=1}^{\tilde{K}} \tilde{\mathbf{a}}_k \circ \tilde{\mathbf{b}}_k \circ \tilde{\mathbf{c}}_k \end{aligned} \quad (15)$$

where  $\mathbf{z}$  is a  $\tilde{K}$ -dimensional vector with its  $k$ th entry given by

$$z_k \triangleq \|\tilde{\mathbf{a}}_k \circ \tilde{\mathbf{b}}_k \circ \tilde{\mathbf{c}}_k\|_F \quad (16)$$

We see that  $\|\mathbf{z}\|_0$  equals to the number of nonzero rank-one tensor components. Therefore minimizing the  $\ell_0$ -norm of  $\mathbf{z}$  is equivalent to minimizing the rank of the tensor  $\mathcal{X}$ .

The optimization (15) is an NP-hard problem. Nevertheless, alternative sparsity-promoting functions such as  $\ell_1$ -norm can be used to replace  $\ell_0$ -norm to find a sparse solution of  $\mathbf{z}$  more computationally efficient. Based on the idea of placing sparsity on the rank-one tensor components, a few CP decomposition algorithms were recently proposed via either optimization techniques [11] or probabilistic model learning [12], [13]. We have no intention to develop a new algorithm in this paper as our objective is to show how to formulate the line spectral estimation problem as a low rank CP decomposition problem. Once an estimate of the factor matrices is obtained, the underlying frequencies can be easily identified since all three factor matrices are Vandermonde matrices and each column of the Vandermonde matrix is associated with an individual frequency parameter.

## V. UNIQUENESS OF CP DECOMPOSITION

Although the factor matrices have a specific structure, we do not need to impose a specific structure on the estimates of the factor matrices since the CP decomposition is unique under very mild conditions. It is well known that essential uniqueness of the CP decomposition can be guaranteed by the Kruskal's condition [14]. Let  $k_{\mathbf{X}}$  denote the k-rank of a matrix  $\mathbf{X}$ , which is defined as the largest value of  $k_{\mathbf{X}}$  such that every subset of  $k_{\mathbf{X}}$  columns of the matrix  $\mathbf{X}$  is linearly independent. Kruskal showed that a CP decomposition  $(\mathbf{A}, \mathbf{B}, \mathbf{C})$  of a three-order tensor is essentially unique if

$$k_{\mathbf{A}} + k_{\mathbf{B}} + k_{\mathbf{C}} \geq 2R + 2 \quad (17)$$

where  $\mathbf{A}, \mathbf{B}, \mathbf{C}$  are factor matrices,  $R$  denotes the CP rank. More formally, we have the following theorem.

**Theorem 1:** Let  $(\mathbf{A}, \mathbf{B}, \mathbf{C})$  be a CP solution which decomposes a three-mode tensor  $\mathcal{X}$  into  $R$  rank-one arrays. Suppose Kruskal's condition (17) holds and we have an alternative CP solution  $(\tilde{\mathbf{A}}, \tilde{\mathbf{B}}, \tilde{\mathbf{C}})$  also decomposing  $\mathcal{X}$  into  $R$  rank-one arrays. Then there holds  $\tilde{\mathbf{A}} = \mathbf{A}\mathbf{\Pi}\mathbf{\Lambda}_a$ ,  $\tilde{\mathbf{B}} = \mathbf{B}\mathbf{\Pi}\mathbf{\Lambda}_b$ , and  $\tilde{\mathbf{C}} = \mathbf{C}\mathbf{\Pi}\mathbf{\Lambda}_c$ , where  $\mathbf{\Pi}$  is a unique permutation matrix and  $\mathbf{\Lambda}_a$ ,  $\mathbf{\Lambda}_b$ , and  $\mathbf{\Lambda}_c$  are unique diagonal matrices such that  $\mathbf{\Lambda}_a\mathbf{\Lambda}_b\mathbf{\Lambda}_c = \mathbf{I}$ .

*Proof:* Please refer to [14]. ■

We now discuss how to choose  $P$  and  $L$  such that the Kruskal's condition can be met. Note that for the line spectral estimation problem, all three factor matrices  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{C}$  are Vandermonde matrices. Hence the k-rank of each factor matrix is equivalent to the minimum value of the numbers of columns and rows, i.e.

$$\begin{aligned} k_{\mathbf{A}} &= \min(N - L - P + 2, K) \\ k_{\mathbf{B}} &= \min(L, K) \\ k_{\mathbf{C}} &= \min(P, K) \end{aligned}$$

In order to satisfy the Kruskal's condition, we can choose one of the three dimensions, say  $P$ , equal to 2, and the other two dimensions  $N - L - P + 2$  and  $L$  no less than  $K$ . Note that when  $R = 1$ , the Kruskal's condition (17) cannot be satisfied whatever  $P$  and  $L$  we choose. Nevertheless, the uniqueness of the CP decomposition also holds for this special case as long as  $\mathcal{X}$  does not contain an identically zero two-dimensional slice along any mode [15].

We would like to emphasize that the tensor  $\mathcal{X}$  is assumed fully observed in Theorem 1 when discussing the uniqueness of the CP decomposition. It still remains an open problem whether the uniqueness holds valid if only a subset of the entries of the third-order tensor are available/observed, which is exactly the situation we are concerned in this paper. This will be a topic of our future investigation.

## VI. SIMULATION RESULTS

We now carry out experiments to illustrate the performance of the proposed method which is referred to as Spectral Compressed Sensing via CP Factorization (SCS-CPF). We compare our method with the Hankel matrix completion method via the projected Wirtinger gradient descent (MC-PWGD) [10], and the atomic norm minimization approach (ANM) [7], [8]. For our method, a Bayesian decomposition technique [13] is employed to perform the CP factorization of the constructed incomplete tensor. The Bayesian algorithm is able to achieve an automatic determination of the rank of the tensor. Also, model parameters associated with the Bayesian approach can be simply chosen and do not require a careful calibration. In our simulations, we choose  $N = 127$  as the MC method [10] requires an odd number of  $N$  to form an  $(N+1)/2 \times (N+1)/2$  low rank Hankel matrix. The frequencies  $\{\omega_k\}$  are uniformly distributed over  $[0, 2\pi)$  and the amplitudes  $\{a_k\}$  are randomly generated according to a normal distribution. The parameters  $P$  and  $L$  are chosen to be 2 and 63, respectively, for our proposed method. Thus  $\mathcal{X}$  is of size  $64 \times 63 \times 2$ . Note that for the noisy case, the ANM method requires the knowledge of the noise variance, which is assumed perfectly known to the

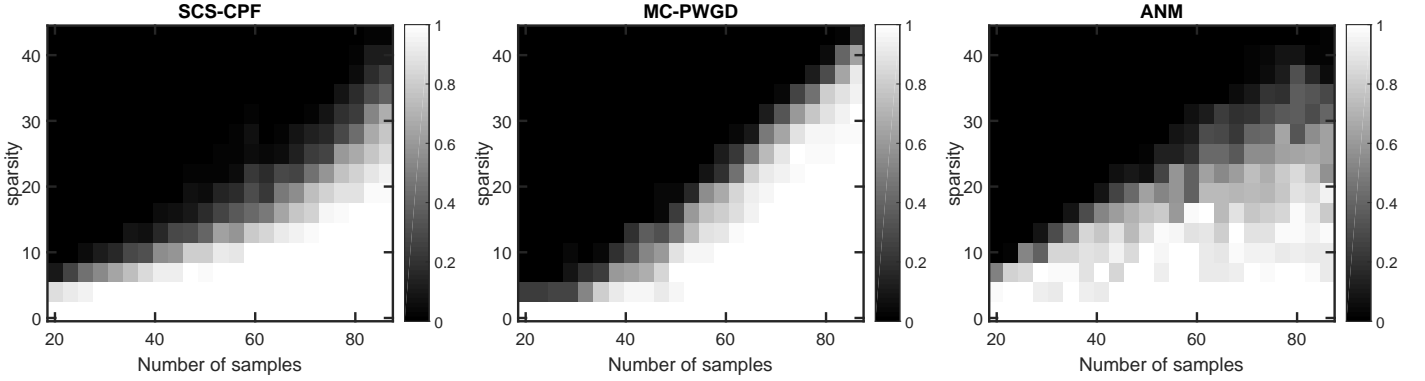


Fig. 1. Phase transitions of respective algorithms.

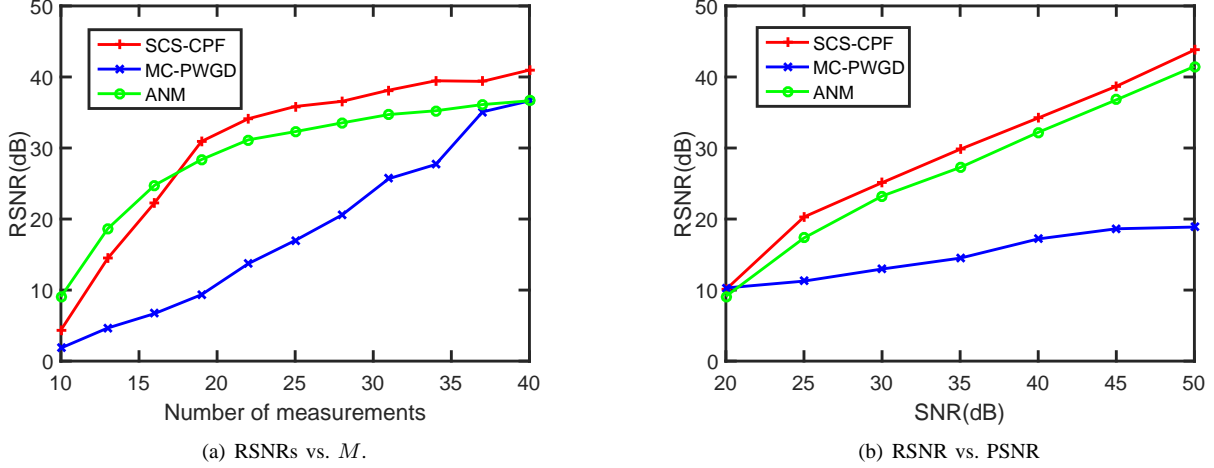


Fig. 2. RSNRs of respective algorithms vs.  $M$  and PSNR.

ANM. Also, the MC-PWGD method requires the knowledge of the number of frequency components, which is assumed known to the MC-PWGD.

We first consider a noiseless case and plot the phase transition curve for each algorithm. We vary the sparsity level  $K = 3 : 2 : 43$  and the number of measurements  $M = 20 : 3 : 86$ . For each point  $(M, K)$ , we conduct 100 independent trials and compute the success rate. A trial is considered successful if the normalized reconstruction error is smaller than  $10^{-3}$ , i.e.  $\|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2 < 10^{-3}$ , where  $\mathbf{x} \triangleq [x_1 \ x_2 \ \dots \ x_N]^T$  denotes the original signal and  $\hat{\mathbf{x}}$  denotes the estimated one. In the phase transition plot, the grey value of each point represents the success rate, with white corresponding to perfect recovery while black corresponding to complete failure. We can see from Fig. 1 that the proposed SCS-CPF method outperforms the MC-PWGD method for a small  $M$  (e.g.  $M \leq 50$ ), where data acquisition is more beneficial due to high compression rates. The ANM method has a sharper transition boundary that the other two methods. Nevertheless, its transition boundary is highly blurred and the size of the white area below the transition boundary is smaller than those of the other two methods, which implies that the ANM is inferior to the other two methods in terms of perfect recovery rates.

We now evaluate the recovery performance of respective algorithms in the presence of additive Gaussian noise. Fig. 2(a) depicts the reconstruction accuracy as a function of signal-to-noise ratio (SNR), where we set  $K = 3$  and  $M = 25$  in our experiments. The reconstruction accuracy is measured by the reconstruction signal-to-noise ratio (RSNR) which is defined as

$$\text{RSNR} = 20 \log_{10} \left( \frac{\|\mathbf{x}\|_2}{\|\mathbf{x} - \hat{\mathbf{x}}\|_2} \right)$$

The reconstruction accuracy as a function of the number of measurements  $M$  is plotted in Fig. 2(b), where we set  $K = 3$  and  $\text{SNR} = 40\text{dB}$ . It can be observed that the proposed method achieves performance similar to the ANM and is more robust against noise than the MC-PWGD method.

## VII. CONCLUSIONS

The line spectral estimation was studied in this paper. We proposed a new method which formulates the observed data into a structured tensor and casts the line spectral estimation problem as a CP decomposition of incomplete tensors. The underlying frequency components can be easily identified from the estimated factor matrices. Simulation results showed that the proposed method provides competitive recovery performance compared with existing state-of-the-art algorithms.

## REFERENCES

- [1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas and Propagation*, no. 3, pp. 276–280, Mar. 1986.
- [2] R. Roy and T. Kailath, "ESPRIT—estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, July 1989.
- [3] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 5, pp. 814–824, May 1990.
- [4] L. Hu, Z. Shi, J. Zhou, and Q. Fu, "Compressed sensing of complex sinusoids: An approach based on dictionary refinement," *IEEE Trans. Signal Processing*, vol. 60, no. 7, pp. 3809–3822, 2012.
- [5] Z. Yang, L. Xie, and C. Zhang, "Off-grid direction of arrival estimation using sparse Bayesian inference," *IEEE Trans. Signal Processing*, vol. 61, no. 1, pp. 38–42, Jan. 2013.
- [6] E. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, June 2014.
- [7] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Information Theory*, vol. 59, no. 11, pp. 7465 – 7490, Nov. 2013.
- [8] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Trans. Signal Processing*, vol. 61, no. 23, pp. 5987–5999, Dec. 2013.
- [9] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Trans. Information Theory*, vol. 60, no. 10, pp. 6576–6601, Oct. 2014.
- [10] J.-F. Cai, S. Liu, and W. Xu, "A fast algorithm for reconstruction of spectrally sparse signals in super-resolution," *Proc. of SPIE, Wavelets and Sparsity XVI*, vol. 9597, Aug. 2015.
- [11] J. A. G. Mateos and G. B. Giannakis, "Rank regularization and Bayesian inference for tensor completion and extrapolation," *IEEE Transactions on Signal Processing*, no. 22, pp. 5689–5703, Nov. 2013.
- [12] P. Rai, Y. Wang, S. Guo, G. Chen, D. Dunson, and L. Carin, "Scalable Bayesian low-rank decomposition of incomplete multiway tensors," in *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, 2014, pp. 1800–1808.
- [13] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 9, pp. 1751–1763, Jan. 2015.
- [14] A. Stegeman and N. D. Sidiropoulos, "On Kruskal's uniqueness condition for the Candecomp/Parafac decomposition," *Linear Algebra and its Applications*, pp. 540–552, 2007.
- [15] N. D. Sidiropoulos, G. B. Giannakis, and R. Bro, "Blind PARAFAC receivers for DS-CDMA systems," *IEEE Trans. Signal Processing*, no. 3, pp. 810–823, Mar. 2000.